A Rate Mapping Template for Alberta Regional Health Authorities

Geographic Methodology Series No. 3

Prepared by
Health Surveillance Branch
Alberta Health and Wellness
Edmonton, Alberta

Latest Revision: February 2004

Alberta
HEALTH AND WELLNESS

For more information contact:

Health Surveillance
Alberta Health and Wellness
P.O. Box 1360
10025 Jasper Avenue
Edmonton, Alberta
T5J 2P4

Phone: 780-422-4518
Fax: 780-427-1470

Series title: Alberta Health and Wellness Geographic Methodology Series
ISBN: 0-7785-2712-3 (online)
Geographic Methodology Series No. 3
A Rate Mapping Template for Alberta Regional Health Authorities
ISBN: 0-7785-3451-0 (pdf)

Note: Users of the information presented in this document in the analysis of health data must insure that products conform to the Alberta *Health Information Act*.

Executive Summary

This document is part of a family of reports that illustrates and documents the geographic methods required to properly analyze health data in Alberta. The descriptions and methods used are consistent across these reports. Together they provide all needed information required to properly understand the spatial component of health data.

This report documents the mapping template created within AH&W Surveillance to properly map rate information. This template allows the user to create consistent maps without the need to learn GIS or other tools. The template is written in Visual Basic and resides embedded in an Excel spreadsheet. No additional plug-ins are required thus avoiding some of the restrictions commonly placed by centrally controlled IT support departments.

The template presents the user's health event data as rates in the form of graphs (rate and standard error), maps, and cartograms (population-based maps).

Acknowledgments

The following team prepared this report:

Alberta Health and Wellness

Erik Ellehoj	Geographic Information System Consultant
Dr. Donald Schopflocher	Biostatistician

The team acknowledges the contributions of the following reviewers in helping to insure completeness and accuracy in the report:

Larry Svenson	Alberta Health and Wellness
Ken Morrison	Alberta Health and Wellness
Joy Edwards	Capital Regional Health Authority

Table of Contents

Executive Summary 3

Acknowledgments 4

Table of Contents 5

Table of Figures 6

I. Introduction 7

II. Enhanced Mapping Methodology 7

III. Automating Mapping Methodology 9

IV. Using the Template 10

V. Subregional Template 15

VI. Conclusion 23

Table of Figures

Figure 1: Rate and Standard Error 9
Figure 2: Macro warning 11
Figure 3: Data worksheet 12
Figure 4: Map worksheet 13
Figure 5: Chart worksheet 14
Figure 6: SubRHA data worksheet 16
Figure 7: SubRHA ABMap worksheet 17
Figure 8: SubRHA Capital worksheet 18
Figure 9: SubRHA Calgary worksheet 19
Figure 10: SubRHA Legend worksheet 20
Figure 11: SubRHA Legend worksheet alternative names 21
Figure 12: SubRHA Graph worksheet 22

I. Introduction

A great demand for health event maps resulted from the creation of AH&W Surveillance. The creation of the branch took place just after the creation of Regional Health Authorities (RHAs) in Alberta in 1994. There was interest in the examination of spatial-temporal patterns in these health data, since much responsibility had been shifted to geographically independent regions. Maps were created for a number of health events using GIS software. A rate-ratio methodology was adopted for all mapping since the rates vary dramatically from one health event to another. The default options for creating area-fill (choropleth) maps in GIS are completely inappropriate for health mapping purposes and yet these types of maps are constantly generated by other jurisdictions. Options such as quantiles for mapping rates result in maps that do not allow any comparisons over time and especially over similar health events.

The rate ratios used were similar to those used in the Cancer Atlas of Alberta. While the rate ratio eliminates many of the problems stated above, it is not able to portray the standard error of the region being mapped. The populations of the regions were dramatically different, therefore the stability of the rate are portrayed in maps and graphs. The problem was temporarily addressed by using patterns to indicate the number of cases within the region. The patterns used increased in ink-coverage as the number of cases increased. The lightest pattern was used when there was only 1 to 2 cases within the region. A denser pattern was used to indicate 3 to 5 cases, denser yet for 6 to 20 cases, and a 100% fill to indicate that there were more than 20 cases.

Map readers were often confused by the combination of five colours with four patterns as it resulted in a maximum of 20 possible pattern-colour combinations within the map. The method did not account for standard error in any manner and any graphs that were produced only showed the regional rates against the provincial rate.

II. Enhanced Mapping Methodology

A new mapping was developed based on a methodology described by Gentleman. The method was first used in a report entitled "Health Trends in Alberta" and has been used for most of the maps portraying rates in all reports generated by AH&W Surveillance.

This method was developed in order to address some issues associated with the variations in population in the Regional Health Authorities (RHA) in Alberta. The population size of a RHA will affect the rate stability associated with that region; rates will be unstable for a small population than for a larger population. The net effect of this methodology is that small variations within a small population will be lessened (or pulled toward the provincial mean) while a small variation with a large population is more important and thus is enhanced (pushed away from the provincial mean. The method listed below is designed to deal with these issues and present the data in a more statistically consistent manner.

The method consists of several steps:

a) Calculate the rates for each region.
b) Calculate the rate for the province.
c) Calculate standard error of a probability of a health event for each regional rate using the following formula:

$$\sqrt{\frac{p(1-p)}{n}}$$

Where: p is the rate (estimate of probability) for the region
n is the number of births for 1 year OR
the summed births for a multi-year calculation
(n is total population for other health events)

d) Calculate the regional specific standard scores. This is accomplished by subtracting the regional rate from the provincial rate and dividing these by the standard score derived for each region in step 3. This is repeated for every region

$$\frac{\text{regional rate} - \text{provincial rate}}{\text{regional standard error}}$$

e) Graph the standard scores calculated in step 4:

Score	Interpretation	Colour
> 2	Significantly Higher than Provincial Average	Red
1 to 2	Higher than Provincial Average	Orange
1 to –1	Average	Yellow
-1 to -2	Lower than Provincial Average	Light Green
< -2	Significantly Lower than Provincial Average	Dark Green

The black dot represents the value of the rate for each region. The colour of the bars above and below the dot represents the score of the region. The portion of the bar closest to the black dot represents the value for a standard score of 1 or –1, while the part of the bars farthest from the dot represent the value for a score of 2 or –2.

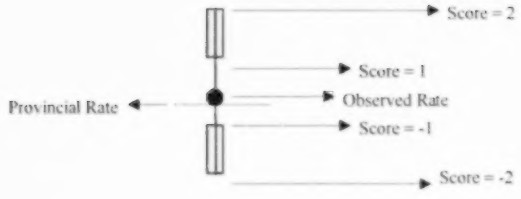


Figure 1: Rate and Standard Error

The figure above illustrates how to interpret the graphic for an individual region. The yellow bars are used to show that provincial rate crosses between the 1 and –1 score range. The table lists other colour possibilities by score category.

f) Generate map using the same categories for each region as listed in step 5.

III. Automating Mapping Methodology

The method described above was well suited to the needs of portraying rates and similar health information for a variety of health events. Unfortunately, this methodology required the knowledge (and access) of sophisticated graphing software along with GIS/mapping software. The resulting graphic images had to be combined within a graphics editing package (such as Corel Draw) in order to include these in a report.

The process was very labour intensive and it discouraged an examination of the data using a variety of combinations since there was a serious time commitment associated with the methodology. A number of options were examined to streamline these processes, including a custom stand-alone program. The use of a stand-alone program provided the greatest flexibility in design, but it also presented problems as most users are not authorized to install any software on their own workstations. Since MS Excel is commonly used for a number of basic analyses and since it is available in the desks of most health analysts it was chosen as platform for building the mapping-graphing template that uses the methodology stated above.

Excel contains a basic mapping module based on the technology on which the MapInfo GIS software was developed. The default installation of Excel does not include this module, and a special installation is required in order to add this module. Since most are not authorized users, or have access to the software disks for this type of installation, this option was discarded.

The mapping functionality was built in Visual Basic and included within the spreadsheet as a macro. The graphing functionality in Excel is very basic. Mapping functionality

was enhanced by using Visual Basic and adapting exiting methods to allow for the sophisticated graphing required.

IV. Using the Template

The template contains five worksheets:

- Data –This worksheet contains the data entered by the user and the formulas required to calculate rates, standard error, score, and other measures required for making the map and graph.
- Chart –This worksheet contains the graph that displays the rate for all regions along with the provincial rate, the standard error for each region, and the assigned category for each region.
- Map –This worksheet contain the map and cartogram which show the regions in the colours associated with the score values calculated for each.
- Legend –This worksheet contains a list of names of the regions.
- Help –This worksheet contains basic help on using the worksheet.

All of these worksheets are properly described below.

To open the template you can double click on the icon or open the Excel spreadsheet once Excel is running. In either circumstance, the following warning appears:

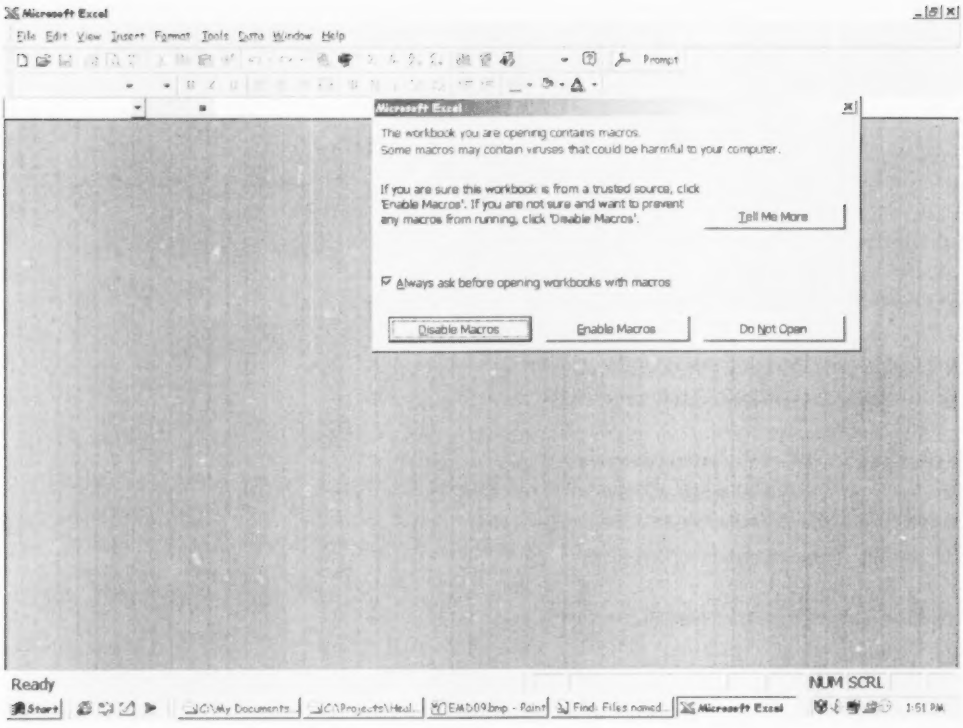

Microsoft Excel
File Edit View Insert Format Tools Data Window Help
Prompt
Microsoft Excel
The workbook you are opening contains macros.
Some macros may contain viruses that could be harmful to your computer.
If you are sure this workbook is from a trusted source, click 'Enable Macros'. If you are not sure and want to prevent any macros from running, click 'Disable Macros'.
Tell Me More
Always ask before opening workbooks with macros
Disable Macros
Enable Macros
Do Not Open
Ready
NUM SCRL
Start
C:\My Documents
C:\Projects\Heal...
EM3D9.bmp - Paint
Find: Files named...
Microsoft Excel
1:51 PM

Figure 2: Macro warning

The default settings in Excel warn the user that there is a macro in a spreadsheet in order to avoid hidden macros that may contain viruses. In this case, the macro is required to generate the map, therefore press the "Enable Macros" button. Please note that some installations of Excel require that the user allow macros to run within their spreadsheet. As the file opens, the user is presented with the Data worksheet:

Figure 3: Data worksheet

The user will be required to copy and then paste into the yellow portion of the data worksheet the cases and population information for the health event being examined. The template automatically calculates the rates, standard erorrs, etc. Please note, however, that these rates are crude rates only as the functionality for creating age-sex standardized rates does not exist within this template. If the user has age-sex standardized rates he/she may paste it in the first light green column. Pasting the rates in this column will delete the formulae and replace them with the appropriate rates calculated elsewhere. The case and population data are still required even if the rate was pasted from an external source. If the calculated provincial rate is different than that calculated by the spreadsheet (cell D3), then the user may also paste this rate into cell D3.

Cell A15 can accept a 0 or a 1. A 0 represents the normal colour settings so that a high rate appears in red and a low rate in green. A value of 1 in this cell reverses the settings so that a high value is presented in green and a low value is presented in red. This type of setting can be used when a high rate is considered a positive factor, such as regular physical exercise.

Cell A18 allows the user to set the rate multiplier so that the "Y" column in the graph shows values that are more appropriate. The user may enter "100" to view percentages, "100,000" to view cases per 100,000, and so on.

Those who are interested in how the standard error or score are calculated may examine any cell in column E or F to determine how this is accomplished. The extra columns on the right are used to "fool" Excel into producing a graph beyond its basics capabilities.

The user may press the purple button (Press here to create the graph and map) only once all of the required information has been entered into the spreadsheet. The macro will cause a number of changes on the screen and finally present the following:

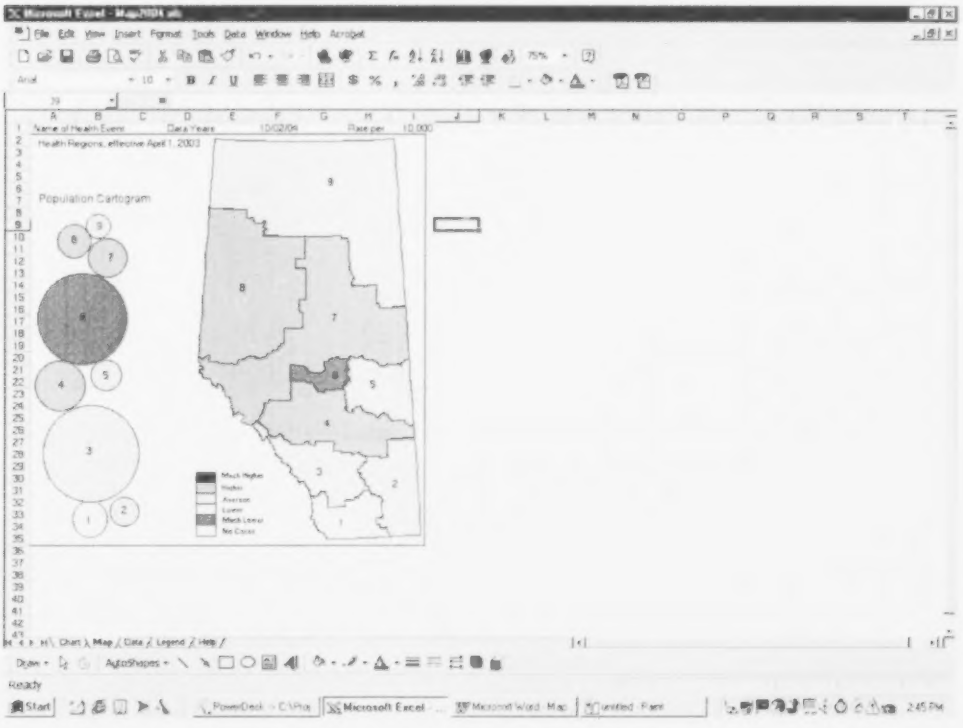


Figure 4: Map worksheet

The worksheet above presents the map and cartogram. Each region is assigned a colour according to their score (calculated in the Data worksheet). A legend explains the colours assignments. When the colour is reversed (for positive factors, such as regular exercise) so do the legend assignments.

The cartogram presents the same information that appears in the map, except that the regions are not drawn based on their geographic extents. The size of each region is proportional to its population. This graphic is useful in determining the population impact of an observed pattern. For example, if the Calgary Health Region and the Capital Regional Health Authority appear in red and the rest of the province in dark green

the impact of the health event appears to be minor. When the same data is mapped using a cartogram, it population impact becomes clear (2/3 of the population).

Cell F1 places the current date for reference. Cell I1 looks up the rate multiplier as it appears in the Data worksheet (cell A18). The user is encouraged to type the name of the health event and the years of data examined in cells A1 and D1, respectively.

To print this worksheet, the normal printing procedure is followed (File/Print). To copy the map to the clipboard, press <Shift> <Ctrl> M followed by Edit/Copy. The user may now paste the map into MS Word, Powerpoint, or other applications. To select the cartogram, press <Shift> <Ctrl> C followed by Edit/Copy. To select the legend, press <Shift> <Ctrl> L followed by Edit/Copy. Please note that all of these objects are MS Office graphic objects and are thus governed by the restrictions posed by that format. These graphic objects have limits on the number of vertices that are much lower than those of a GIS and thus many boundaries have a "choppy" appearance. All of these graphics are quite acceptable for use in presentations and for draft version of a report. Unfortunatelly, for final versions of reports, a traditional method of GIS and graphics illustration package must be used.

The same information also appears in the form of a graph:

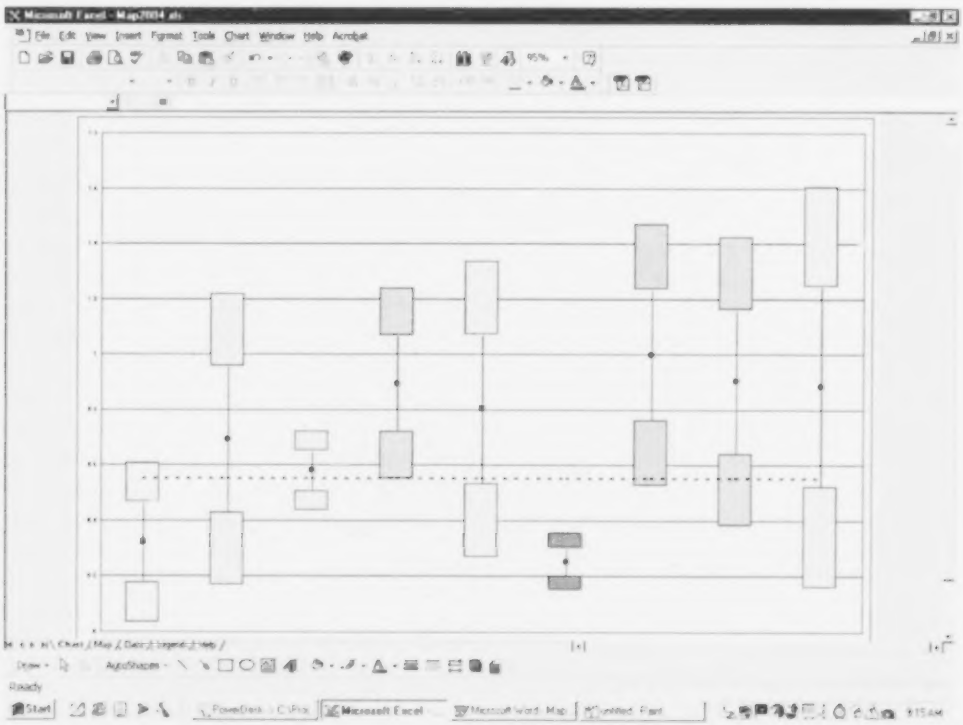


Figure 5: Chart worksheet

This view presents a very different view of the same data. While the Map worksheet illustrated the geographic patterns and population impact, this view describes the regional rate, provincial rate, and regional standard error.

Each vertical graphic elements represents a Regional Health Authority (RHA). The Y axis presents the rate as defined by the rate multiplier in cell A18 in the Data worksheet. The dashed line that runs horizontally through the graph represents the provincial rate. The colours assigned to each vertical graphic element are consistent with those that appear for the same RHA in the map and cartogram views.

The dot in the centre of each vertical graphic element represents the rate for that RHA. The vertical distance separating the dot from the start of the coloured bar represents one standard error for that particular RHA. It can be argued that any rate observed within one standard error of the provincial rate is the same as the provincial rate. For example, the rate fc region 9 (Northern Lights) is much higher than the provincial rate, however since it also reports a large standard error, the reported rate is considered to be in the same category as the provincial rate. The rate for region 8 (Peace Country) is only slightly above the rate for region 9, but since it has a smaller standard error (resulting from a larger population) its category is different than the region 9's category. The vertical distance of each coloured bar also represents one standard error. The rate for region 8 is within 1 to 2 standard errors of the provincial rate (since the dashed line appears over the coloured portion of the vertical bar (which represents 1 to 2 standard errors for that region). Region 1 is also between 1 and two standard errors of the provincial rate, but in this case below the provincial rate. If the provincial rate rate is not within one or two standard errors, then it is placed into a separate category that highlights that the rate is significantly (but not necessarily statistically) different from the provincial mean. In all likelihood, most of these instances will be statistically significant from the provincial mean. Region 6 is significantly lower than the provincial mean as its regional mean is more than two standard errors from the provincial rate.

V. Subregional Template

The template described above was originally constructed for the 17 RHAs that existed prior to 2003. As a result of the amalgamation of these regions, only 9 regions are now available for reporting health data. As these regions have become larger, the regional differences have decreased as the new regions have larger physical area and population. A set of subregional reporting units were created in conjunction with the new 9 RHAs, and these units are described in a document entitled "Small Area Analysis: Definition of Sub-RHA Geographic Units in Alberta". The resulting 68 subregions have populations large enough to allow for stable rate calculations for a large proportion of health events. Information about rates necessary for mapping at this level appear later in this document.

The success of this methodology lead to requests for maps reporting to this level of information. A postal code to subRHA lookup file was created in order to allow summarizing of any health event data to these boundaries. An EA to subRHA lookup file

was also created in order to allow census data to be summarized at the same level of information.

To complement this technology, a subRHA template has been created. At the current time, it is available on request for research purposes only.

The template is very similar to the RHA mapping/graphing template, except that more worksheets are required due to the greater amount of information being portrayed. The first worksheet after opening the spreadsheet (and receiving the warning about macros).

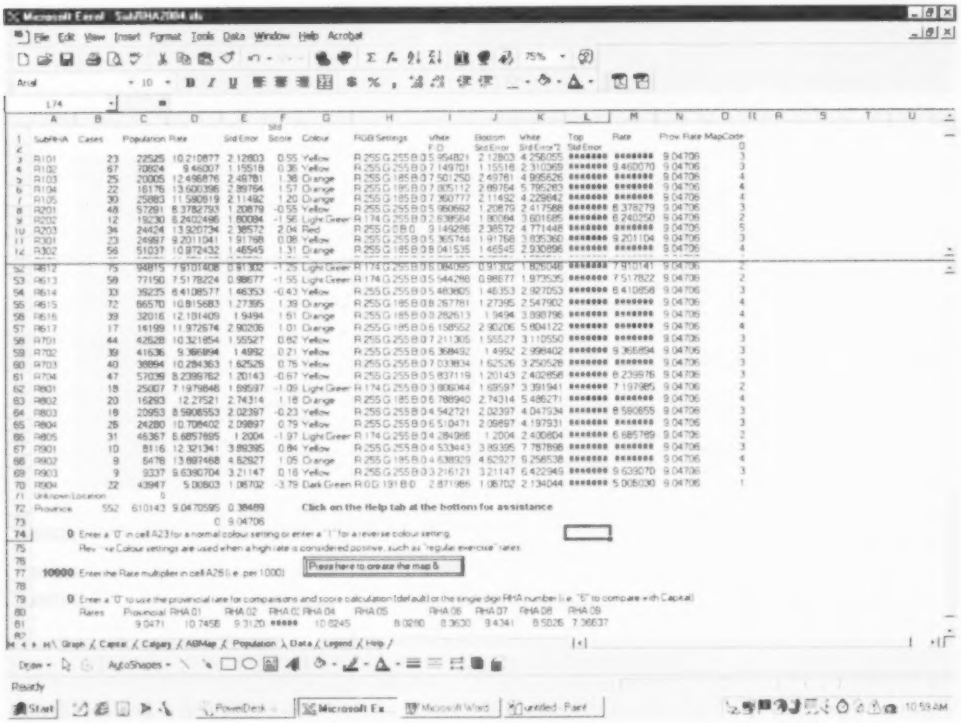


Figure 6: SubRHA data worksheet

This sheet is very similar to the one that appears in figure 4. The worsheet has been split horizontally in order to show the top and bottom of the sheet at the same time. The user may, once again, paste the case and population data into the yellow section of the worksheet. If age-sex standardized rates have been calculated elsewhere, these can be pasted into column D. The scores are calculated automatically.

The colours can be reversed by entering a "1" in cell A74 in order to portray a positive health event, such percentage of people who quit smoking. The same functionality appears in the RHA mapping template. The rate multiplier can be entered into cell A77 and it offers the same functionality as its equivalent in the RHA mapping template.

Cell A79 accepts values from 0 to 9, and these correspond to the base line used for comparisons. A value of "0" is the default and it corresponds to having every subRHA compared against the provincial rate (as is the case in the RHA mapping template). Other options are available for individual RHAs who wish to compare all the subRHAs within the RHA against the rate for the RHA as a whole. A number of health events have a rural-urban pattern which results in most subRHAs from a single RHA appearing in the same category. By comparing the rates of the subRHAs against the respective RHA, it is possible to determine if there are any issues at a level lower than the RHA. A value of "3" will be entered in the cell if RHA 3 (Calgary) wishes to compare the rates of the Calgary subRHAs against the rate for region 3.

Pressing the purple button executes the macros that generate the maps and graphs. This set of macros runs slower than those in the RHA mapping template as it has to process almost ten times as much information. The provincial map appears in the ABMap worksheet.

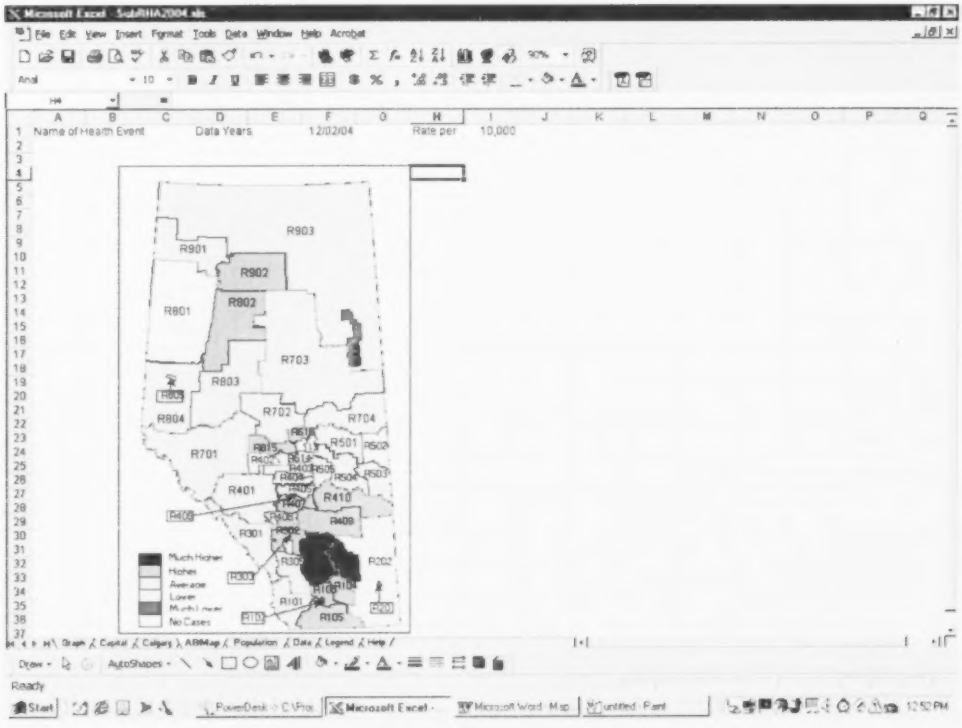


Figure 7: SubRHA ABMap worksheet

In this worksheet each subRHA is assigned a colour according to the classification of the health data entered in the Data worksheet. A blue line was used on top of the subRHAs

to help identify the RHA boundaries. The second digit of the subRHA also identifies the membership RHA for the subRHA.

The user is encouraged to enter the name of the health event portrayed by the map in cell A1, and the data years in cell D1. The current date appears in cell F1 and the rate multiplier in cell I1. The legend appears in the bottom left corner of the map.

To select the entire map, press <Shift> <Ctrl> M and then Edit/Copy. The graphic in the clipboard can then be pasted into any application, such as MS Word or Powerpoint. To select the legend, press <Shift> <Ctrl> L and then Edit/Copy. To select the subRHAs for any particular RHA, press <Shift><Ctrl># where # corresponds to the RHA number (i.e. 5 for East Central), followed by Edit/Copy.

To print the worksheet, simply select File/Print, the entire map fits in an 8.5 x 11" page.

The subRHAs in the two largest areas in the province can not be shown in this map as their geographic size is too small to be discerned at a provincial level. Two separate worksheets were created to accommodate these two urban areas in detail.

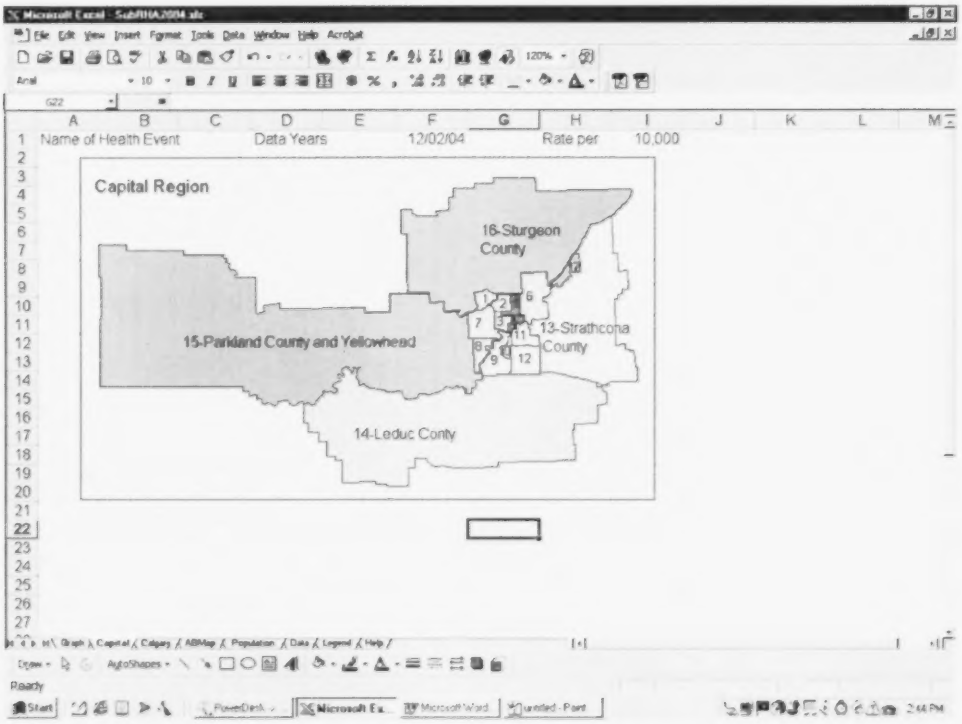


Figure 8: SubRHA Capital worksheet

This worksheet is very similar to the AB Map worksheet, except that only the Capital (6) RHA is shown in detail. To select all of this map, press <Shift> <Ctrl>6 and then Edit/Copy. The map may then be pasted in any application.

A similar map is available for the Calgary region.

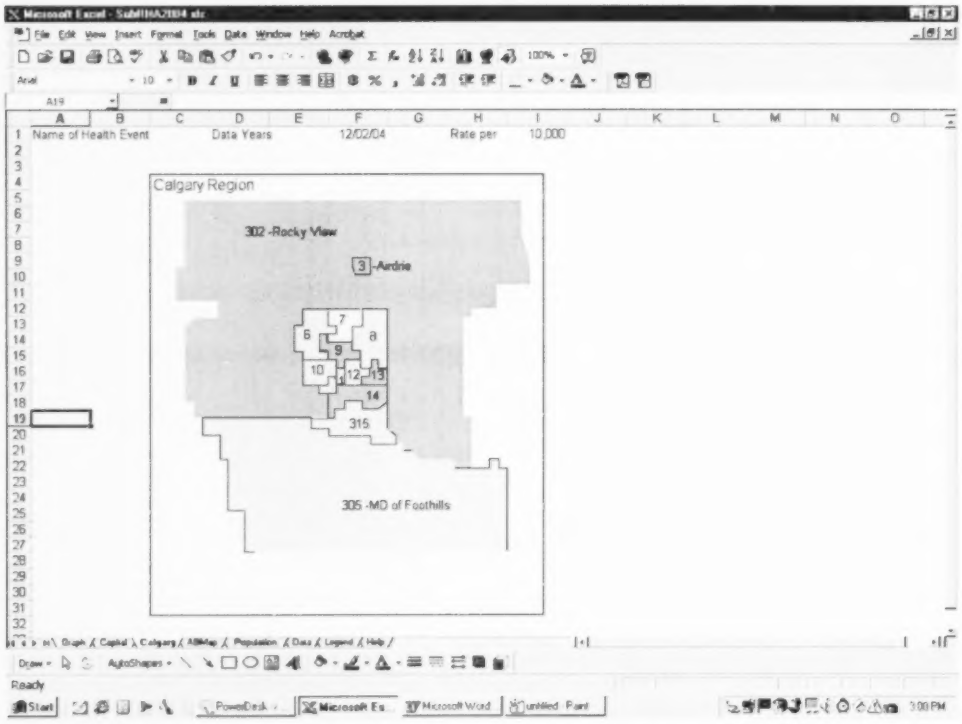


Figure 9: SubRHA Calgary worksheet

This map is similar to the one for the Capital region, with one exception; it does not show the entire RHA as this region is now quite large. It would be impossible to see the subRHAs in the City of Calgary if the entire region was portrayed. Subregions R301 (Banff-Kananaskis) and R304 (Wheatland) are the only missing components, while R302 (Rocky View) and R305 (MD of Foothills) are shown only partially.

A worksheet entitled "Legend" contains a list of all the subRHAs and an assigned name. These names may be revised as these boundaries are used more commonly by the individual RHAs.

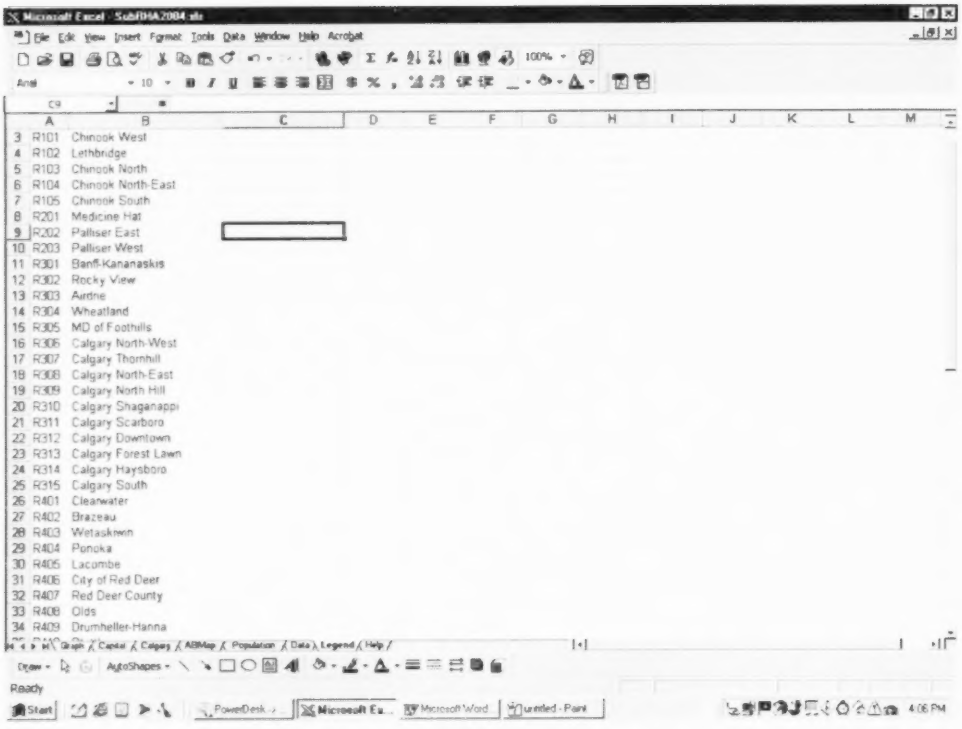
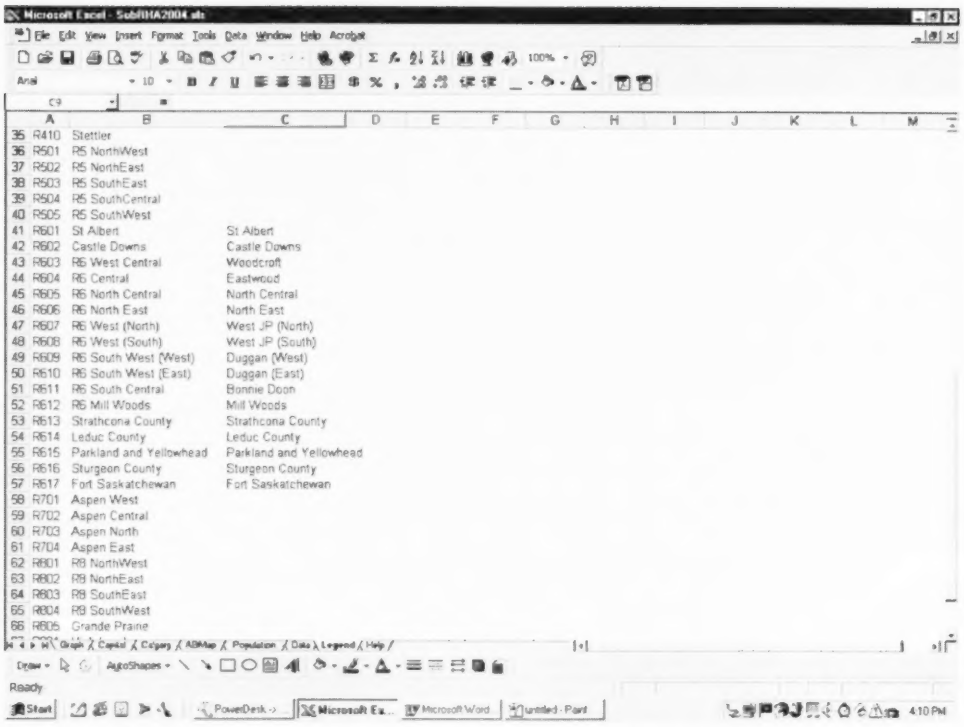

Microsoft Excel - SubRHA2004.xls
File Edit View Insert Format Tools Data Window Help Acrobat
A B C D E F G H I J K L M
3 R101 Chinook West
4 R102 Lethbridge
5 R103 Chinook North
6 R104 Chinook North-East
7 R105 Chinook South
8 R201 Medicine Hat
9 R202 Palliser East
10 R203 Palliser West
11 R301 Banff-Kananaskis
12 R302 Rocky View
13 R303 Airdrie
14 R304 Wheatland
15 R305 MD of Foothills
16 R306 Calgary North-West
17 R307 Calgary Thornhill
18 R308 Calgary North-East
19 R309 Calgary North Hill
20 R310 Calgary Shaganappi
21 R311 Calgary Scarboro
22 R312 Calgary Downtown
23 R313 Calgary Forest Lawn
24 R314 Calgary Haysboro
25 R315 Calgary South
26 R401 Clearwater
27 R402 Brazeau
28 R403 Wetaskiwin
29 R404 Ponoka
30 R405 Lacombe
31 R406 City of Red Deer
32 R407 Red Deer County
33 R408 Olds
34 R409 Drumheller-Hanna
Ready
Start
4:06 PM

Figure 10: SubRHA Legend worksheet

In some cases alternative names exist for the same subRHA. In these cases both are included, in separate columns.



	A	B	C
35	R410	Stettler	
36	R501	R5 NorthWest	
37	R502	R5 NorthEast	
38	R503	R5 SouthEast	
39	R504	R5 SouthCentral	
40	R505	R5 SouthWest	
41	R601	St Albert	St Albert
42	R602	Castle Downs	Castle Downs
43	R603	R6 West Central	Woodcroft
44	R604	R6 Central	Eastwood
45	R605	R6 North Central	North Central
46	R606	R6 North East	North East
47	R607	R6 West (North)	West JP (North)
48	R608	R6 West (South)	West JP (South)
49	R609	R6 South West (West)	Duggan (West)
50	R610	R6 South West (East)	Duggan (East)
51	R611	R6 South Central	Bonnie Doon
52	R612	R6 Mill Woods	Mill Woods
53	R613	Strathcona County	Strathcona County
54	R614	Leduc County	Leduc County
55	R615	Parkland and Yellowhead	Parkland and Yellowhead
56	R616	Sturgeon County	Sturgeon County
57	R617	Fort Saskatchewan	Fort Saskatchewan
58	R701	Aspen West	
59	R702	Aspen Central	
60	R703	Aspen North	
61	R704	Aspen East	
62	R801	R8 NorthWest	
63	R802	R8 NorthEast	
64	R803	R8 SouthEast	
65	R804	R8 SouthWest	
66	R805	Grande Prairie	

Figure 11: SubRHA Legend worksheet alternative names

The graph to accompany these maps is obviously more complex than the graph used for the RHA template, however the logic is consistent.

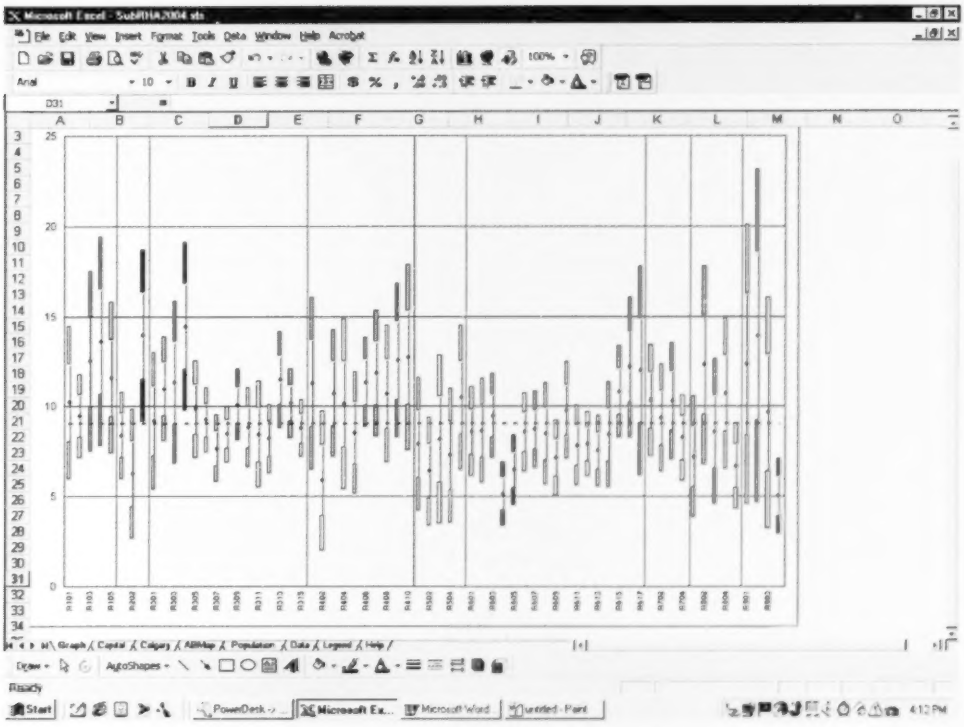


Figure 12: SubRHA Graph worksheet

Once, again, the provincial rate appears as a horizontal dashed line. Note, however, that if the value in cell A79 in the Data worksheet is not "0", then the dashed line will correspond to whatever RHA rate that was selected (1 through 9). As the value in cell A79 in the Data worksheet is changed, the purple button must be pressed again to update the colours in the graph and maps.

The black dots represent the observed rate for each subRHA. The vertical distance from the dot to the start of the bar reprents one standard error for that subRHA. The distance between the start and the end of the bar represents another standard error for the same subRHA. The categories and colour assignments are the same as those used for the RHA mapping template. The observed standard errors are larger in this graph, however, as this is the logical outcome from smaller populations.

The labels for each subRHA appear for every second subRHA due to space restrictions. Please use the Legend worksheet to determine the name of each subRHA or examine the map to determine its geographic location.

The rates that appear on the left side (Y axis) correspond to the calculated (or pasted) rates after conversion by the multiplying factor in cell A77 in the Data worksheet.

VI. Conclusion

The authors welcome any suggestions about improvements to this document and to the mapping/graphing Excel templates. An effort has been made to provide as much functionality as possible, while ensuring that these tools are simple to use and require as little user-intervention as possible. Any maps that will be used for publication should be generated using a mapping or GIS program followed by a graphics illustration program. These tools allow for quick analysis of any health data at the RHA and SubRHA level without the need for expensive tools. Draft versions of reports can be easily generated using this tool and a more sophisticated solution is only for final publication.

References

Ellehoj, E.A., Schopflocher, D. et al (2004). *Calculating Demographic and Epidemiological Quantities in Alberta by Geo-Political Areas Boundaries-Geographic Methodology Series No. 1.*

Ellehoj, E.A., Schopflocher, D. et al (2004). *Using the AH&W Lookup File to Aggregate Data to Different Geographic Boundaries-Geographic Methodology Series No. 2.*

Ellehoj, E.A., Schopflocher, D. et al (2004). *Calculating Distances in Alberta-Geographic Methodology Series No. 4.*

Ellehoj, E.A., Wilson, S. et al (2005). *Definition of Rural and Urban Areas for Health Services Provision-Geographic Methodology Series (in preparation).*

Ellehoj, E.A., Schopflocher, D. et al (2005). *Geographic Activities in AH&W - Geographic Methodology Series (in preparation)..*

Health Trends in Alberta, 2001. Alberta Health Surveillance, Alberta Health and Wellness. ISSN 1480-6657